

特集にあたって

後藤 順哉 (中央大学)

近年、深層学習に代表される機械学習、信号処理、画像処理などのさまざまな分野で、一次法 (first-order methods) と総称される最適化アルゴリズムが重要な役割を果たしている。本特集はその背景や理論の一端を紹介するものである¹。

過去 10 年の間に最適化関連の国際会議でも一次法に関するセッションが存在感を増し、そのみを扱った出版物 ([1] など) や解説記事も多く目にするようになった。しかし、その基本的アルゴリズムはすでに 1960~80 年代に提案されているが、実用的な観点からは長いことあまり注目されてこなかったものが多い。実際、90 年代後半に最適化を学んだ筆者にとって、一次法と言えば最急降下法がその代表であり、講義ではその後に扱われるニュートン法や準ニュートン法の、言わば「かませ犬」的な位置づけという存在であった。「一次収束しかない」「無駄にジグザグする」といった印象を植え付けられつつ、二次収束はかなり難しいの間を取って「超一次収束の手法が落としどころか」という気分が陥るといった具合である。それが今や、上述の分野で重宝されている一次法は普通には一次収束すらせず、加速する方法を工夫して一次収束できたら 100 点満点という。そして二次法に至っては上述の分野において実用的でないとのこと。

80 年代に線形計画 (LP) に革新的進歩をもたらした内点法はニュートン法をベースにした典型的な二次法である。内点法登場のインパクトは大きく、90 年代には LP ソルバーの飛躍的進歩や二次錐計画 (SOCP)、半正定値計画 (SDP) といった問題クラスの拡大が最適化研究のメインストリームをなした感があり、日本の OR とりわけ連続最適化コミュニティもその流れの中で大きな貢献をなした。時が経ち 2000 年代半ばくらいに状況が変わり出す。冒頭に挙げたデータ分析を扱う分野で一次法の一斉蜂起が始まる。それまで毎年増加を続けた CPU のクロック周波数の成長は頭打ち

になり、マルチコア CPU、並列計算といった言葉を多く聞くようになり、計算を巡る状況も変わったようにも思えた。そうこうしているうちに人工ニューラルネットワークが深層学習として生き返る…。最適化に関して言えば、2000 年代半ばからの 10 年はそれまで見向きもされなくなっていた技術が次々と蘇りメインストリームに返り咲いた 10 年に映る。本特集の題目にした「逆襲」はそのような状況の比喩である。

そんな背景もあり、現在の一次法の興隆は冒頭に挙げた分野を専門とした研究者や一次法に対する偏見 (あるいはトラウマ) をもたない世代が牽引してきた印象が強い。しかし最適化は最適化である。二次法もそうであったように、一次法にも長所と短所/可能性と限界があり、いろいろな見方・立場からの新しい展開も期待できるのではないだろうか。とはいえ、機械学習などの分野ではカンファレンスペーパー駆動で猛烈な勢いで研究が進んでおり、キャッチアップするのは容易でない。貢献していくのであれば、研究の先端がどこにあるのか、どこに向かっているのか、さまざまな概念、その歴史、それらの関係、つまり文脈を踏まえることが有効であろう。それを知る一助になるのが本特集の狙いである。

そこで、当該分野で研究を進めている研究者諸氏にご協力いただき、各自の最近の研究の紹介をベースに、周辺の技術の紹介をしていただくこととした。

さて、一口に一次法と言っても、それが指す範囲には若干ぶれがあるようである²。本特集では大雑把に「すべてまたは一部の変数に関する勾配情報 (勾配ベクトルや劣勾配) のみを用いる反復解法」を指すことにする。

¹ 特集の題目がいくぶん映画「スターウォーズ」風なのは今年同作最終作が公開されることと無縁ではない。ちなみに映画で反乱軍が対峙する敵軍は「ファーストオーダー」を名乗っている。ただし本特集の内容は映画とは全く無関係である。

² 山下氏による本特集記事によれば、あくまでも一次法は勾配ベクトルを用いるもので、劣勾配を用いる場合の近接勾配法などは一次でなくゼロ次法とのことである。また単に最適化問題を記述する関数の一次の情報 (勾配ベクトルなど傾き) を使うか二次の情報 (ヘッセ行列など曲率) を使うかだけでなく、高々一次収束しかないのが一次法という見方も提示されている。 p 次収束の定義については専門書か成島氏による記事の脚注を参照いただきたい。

ごとう じゅんや

中央大学理工学部経営システム工学科

〒 112-8551 東京都文京区春日 1-13-27

jpgoto@indsys.chuo-u.ac.jp

逆に、関数の二次の情報であるヘッセ行列を直接扱うニュートン法、二次関数により近似する逐次二次計画法 (SQP)、信頼領域法は含まない。一方で、準ニュートン法と呼ばれるものにはいくつかバリエーションがあり、ヘッセ行列を近似した行列そのものを更新するもの (たとえば BFGS 公式によるもの) は含めず、近似行列を勾配ベクトルのみから構成するような記憶制限法などは含める³。

はじめの三つの記事は勾配ベクトルが使えない問題にも適用可能なアルゴリズムを扱う。

小野峻佑氏 (東京工業大学) による一つ目の記事は、画像/信号処理分野においてなぜ一次法が発展したのかに始まり、三つの近接分離アルゴリズム (近接勾配法、交互方向乗数法、主-双対近接分離法) が具体的にどのようなアルゴリズムで、それぞれどのような問題に適用され成功してきたかを紹介している。全体的に技術的ながら、一つひとつの専門用語の定義が過不足なく丁寧に記述されているので、各アルゴリズムの要点を理解するのにも役立つだろう。

伊藤勝氏 (日本大学) の記事は小野氏の記事で紹介されたアルゴリズムを含む一次法の一般的な枠組みを俯瞰し、問題がもつ性質に対応して一次法の限界に関する既知のさまざまな理論的結果を一つの流れとして見事にまとめ上げている。内容はやや技術的に高度なため、初学者にはやや難しいかもしれないが、最適化アルゴリズムの基礎を知っているが最近の一次法について部分的にしか知らないという専門家には一読の価値がある記事となっている。

小野氏、伊藤氏が若手の、一次法が不遇を買っていた時代を知らない世代であるのに対し、3番目の記事の執筆者、山下信雄氏 (京都大学) は筆者と同じような場面を見てきた (少し上の) 世代に属しつつ、一次法も二次法も研究対象とし、最適化アルゴリズム研究の王道を突き進んでこられた立場からの非常に興味深い見方を記事の中で展開している。また、一次法が活躍する問題の特殊構造やいろいろな工夫など具体的なヒントに溢れた、ためになる1編である。ADMMや準ニュートン法など小野氏や (前後するが) 成島氏の記事に書かれたアルゴリズムを頭に入れてから読むのが、

お勧めの1編である。

続く成島康史氏 (慶応義塾大学) の記事は、勾配ベクトルが存在する場合の、掛け値なしの一次法を扱ったものである。特に (フルランクの) 行列形式でデータを保持しないようにしながら最急降下法を改良する方法として、記憶制限準ニュートン法と非線形共役勾配法の2タイプの方法のエッセンスと、両者の関係について丁寧に説明していただいた。これらの方法は1反復の計算手間を抑えながら一次法を改良していく素直な方向を指し示しているとも言え、今後も見据えながら読むと新しい展開に繋がるかもしれない。

最後の2編はDCと呼ばれる構造をもった非凸最適化問題を扱っており、凸最適化に焦点を当てたそれまでの4編と対照的である。5本目の筆者らによる記事はスパース最適化のうち、非ゼロ要素数という不連続な関数を一定値以下に抑える制約に対する連続関数を用いた等価表現の紹介と、それを制約に持ったDC最適化問題に対する近接勾配法と同様のアルゴリズムやADMMなどを紹介している。

大規模な問題に対しては勾配計算すらサボることが一つの有効な戦略となりうるが、ランダムに選択した (勾配) 要素を用いる方法である確率勾配法は実用上も理論上も大変ホットな方法論である。当該手法の研究で有名な鈴木大慈氏 (東京大学) らによる6本目の記事では、DC最適化問題に対する最近の研究についてご紹介いただいている。

このように現在の一次法ブームを支えるいくつかの重要なアルゴリズムを扱ってはいるが、含まれていない重要なアルゴリズムもある。それらについては各記事末尾の参考文献リストを参考に、専門書や論文を参照いただきたい。

最後にご協力いただいた各執筆者に感謝申し上げます。また、担当編集委員の伊豆永氏 (神奈川大学) には本特集のとりまとめをご依頼いただき、原稿のチェックなどご協力いただいた。記して感謝したい。

参考文献

- [1] A. Beck, *First-Order Methods in Optimization*, MOS-SIAM Series on Optimization, 2017.

³ ちなみに LP の単体法はまさに一次法である。しかし上記の文脈を踏まえると少し趣旨がずれるし、前々号の4月号で豪華執筆陣による「はじめよう線形計画法」が企画されたこともあり、本特集には含めていない。